MULTIPLE AND DISSOCIABLE EFFECTS OF SENSORY HISTORY ON WORKING-MEMORY PERFORMANCE

A PREPRINT

Jasper E. Hajonides
Oxford Centre for Human Brain Activity,
University of Oxford, UK
Department of Experimental Psychology,
University of Oxford, UK

Freek van Ede
Department of Applied and Experimental Psychology,
Vrije Universiteit, NL

Mark G. Stokes
Department of Experimental Psychology,
University of Oxford, UK

Anna C. Nobre
Oxford Centre for Human Brain Activity,
University of Oxford, UK
Department of Experimental Psychology,
University of Oxford, UK

Nicholas E. Myers*
School of Psychology
University of Nottingham, UK
Department of Experimental Psychology,
University of Oxford, UK

October 31, 2021

ABSTRACT

Behavioural reports of sensory information are biased by stimulus history. The nature and direction of such serial-dependence biases can differ between experimental settings – both attractive and repulsive biases towards previous stimuli have been observed. How and when these biases arise in the human brain remains largely unexplored. They could occur either via a change in sensory processing itself, post-perceptual maintenance or decision-making processes, or both. Here, we analysed behavioural and magnetoencephalographic data from a working-memory task in which participants were sequentially presented with two randomly oriented gratings, one of which was cued for recall at the end of the trial. Behavioural responses showed evidence for two distinct biases: 1) a within-trial repulsive bias away from the previously encoded orientation on the same trial, and 2) a between-trial attractive bias towards the task-relevant orientation on the previous trial. Multivariate classification of stimulus orientation revealed that neural representations during stimulus encoding were biased away from the previous grating orientation, regardless of whether we considered the within- or between-trial prior orientation – despite opposite effects on behaviour. These results suggest that repulsive biases occur at the level of sensory processing and can be overturned at post-perceptual stages to result in attractive biases in behaviour.

Keywords Working Memory · Attention · Serial Dependence · Neural Representations · Supervised Learning · Perceptual Bias

*Corresponding author. E-mail address: Nick.Myers@nottingham.ac.uk

1 Introduction

Stimulus history modulates performance based on sensory input. The incorporation of recently perceived information is a beneficial strategy within a world that is largely stable over short time scales (Dong & Atick, 1995). This reliance on temporal correlations is deeply engrained in the visual system (Simoncelli & Olshausen, 2001). Leveraging past sensory evidence can be an effective prior to extract signals from the noisy sensory stream and to help maintain a stable representation to bridge blinks, eye movements, or visual occlusions.

Recent studies have revealed robust short-term influences of prior stimulation in delayed-response and working-memory tasks (Bae & Luck, 2017; Cicchini, Anobile, & Burr, 2014; Cicchini, Mikellidou, & Burr, 2018; Czoschke, Fischer, Beitner, Kaiser, & Bledowski, 2019; Czoschke, Peters, Rahm, Kaiser, & Bledowski, 2020; Fischer & Whitney, 2014; Fritsche, Mostert, & de Lange, 2017). In these studies, perceptual decision making is systematically influenced by stimulus features presented earlier in the trial (Bae & Luck, 2017; Czoschke, Fischer, Beitner, Kaiser, & Bledowski, 2019; Czoschke, Peters, Rahm, Kaiser, & Bledowski, 2020; Fritsche & de Lange, 2019; Fritsche, Mostert, & de Lange, 2017), in the previous trial (Cicchini, Mikellidou, & Burr, 2017; Cicchini, Mikellidou, & Burr, 2018; Fritsche, Mostert, & de Lange, 2017; Makovski & Jiang, 2008), or multiple trials back (Fritsche, Solomon, & De Lange, 2021; Fritsche, Spaak, & de Lange, 2020; Gekas, McDermott, & Mamassian, 2019; Suárez-Pinilla, Seth, & Roseboom, 2018). More specifically, items are reported as more *similar* to the task-relevant stimulus on a previous trial. This attractive bias, sometimes referred to as the *serial-dependency bias* (Cicchini, Anobile, & Burr, 2014; Fischer & Whitney, 2014), has been suggested to induce temporal stability in our perceptual image by acting as a prior that systematically biases present perceptual information towards recent stimuli (Fritsche, Spaak, & de Lange, 2020; Kiyonaga, Scimeca, Bliss, & Whitney, 2017). Conversely, item features presented in sequence within trials are often judged more *dissimilar* to each other (Born & Tootell, 1992; Fritsche, Mostert, & de Lange, 2017; Störmer & Alvarez, 2014). The magnitude of this repulsive bias varies widely with experimental design, contrast, presentation duration, and stimulus size, lasting tens of milliseconds up to multiple seconds (Fritsche, Solomon, & De Lange, 2021; Fritsche, Spaak, & de Lange, 2020; Patterson, Wissig, & Kohn, 2013; Priebe & Lisberger, 2002; Suárez-Pinilla, Seth, & Roseboom, 2018). Within this context, the observed biases could be adaptive for exploiting temporal redundancy and optimising perceptual decision-making (Cicchini, Mikellidou, & Burr, 2018; Kiyonaga, Scimeca, Bliss, & Whitney, 2017; van Bergen & Jehee, 2019).

Despite their opposite directions, attractive and repulsive performance biases have been shown jointly to influence task processing (Czoschke, Fischer, Beitner, Kaiser, & Bledowski, 2019; Fritsche, Mostert, & de Lange, 2017; Fritsche, Spaak, & de Lange, 2020; Sadil, Cowell, & Huber, 2021). The use of continuous responses in such working-memory tasks provides an overt means of measuring the feature representation maintained in memory with high sensitivity and can detect small but systematic response offsets (Fisher & Whitney, 2014; Huang & Sekuler, 2010). Yet, little is known about the neural mechanisms of performance biases. From behavioural studies alone, it can be difficult to infer what stage(s) of processing are modulated (since performance measures only provide evidence about the final state of the system, but see Kim et al. 2020).

The current study combined a precision working-memory task for orientation recall and magnetoencephalographic (MEG) recordings. In each trial, participants viewed two orientations and were cued to reproduce one of them using a continuous response. We investigated the neural dynamics of

performance biases resulting from stimulus sequences between and within trials. In addition, we manipulated the task-relevance of grating orientations within trials in order to probe the degree of automaticity of performance biases.

We hypothesised that, in line with Czoschke et al. (2019), we would observe both within-trial and between-trial performance biases. To identify the neural correlates of both types of bias we trained a classifier on MEG sensor data and estimated if the evidence for the presented feature was modulated by stimuli presented on the same trial or on the previous trial. If within-trial repulsive performance biases originate from modulation of sensory stages of processing, this would be reflected in the classification results. Furthermore, if this repulsive bias is an automatic modulation and not under top-down control, the shift in the representation would be the same regardless of the task relevance of the stimulus feature for the subsequent recall task. Similarly, if the attractive between-trial bias originates at the sensory processing stage, a shift in the neural representation would be expected to occur towards the target orientation on the previous trial. Based on previous research (Bae & Luck, 2020; Fischer, Czoschke, et al., 2020), we expected that between-trial biases would be modulated by the task-relevance of the item on the previous trial, such that shifts in representation would be restricted to or significantly larger for previously task-relevant items.

Previewing the results, we confirmed that behavioural responses were systematically repulsed away from previous orientations on the same trial and attracted towards orientations recalled on the previous trial. Mirroring the behavioural repulsion results, the neural representation of stimulus orientation was shifted away from the preceding orientation presented on a given trial. This representational bias was only present when the second item was task-relevant. However, we found no neural evidence of an attractive between-trial bias during orientation encoding. Instead, there was a repulsive neural bias away from the target item on the previous trial. This could be caused by lingering sensory information from the previous trial. We speculate that the behavioural attractive bias between trials may represent modulation at post-perceptual levels of processing, such as during memory maintenance, decision-making, or response preparation.

2 Results

To study sensory history effects, 20 human observers completed a working-memory task while their neural activity was monitored using MEG (see **Methods**). Participants were sequentially presented with two randomly oriented Gabor gratings. A central colour cue presented together with the second grating indicated which of the two orientations participants had to report via continuous recall at the end of the trial (*protect* condition: recall the first orientation or *update* condition: recall the second orientation; see **Figure 1**). We reasoned that after encoding the first grating, participants responded to the cue by either protecting their memory content from interference of the second grating or updating working memory to represent the second grating orientation. Out of the 400 trials that each participant completed, 200 trials contained two orientations. The other 200 trials served as a 1-item memory-load control.

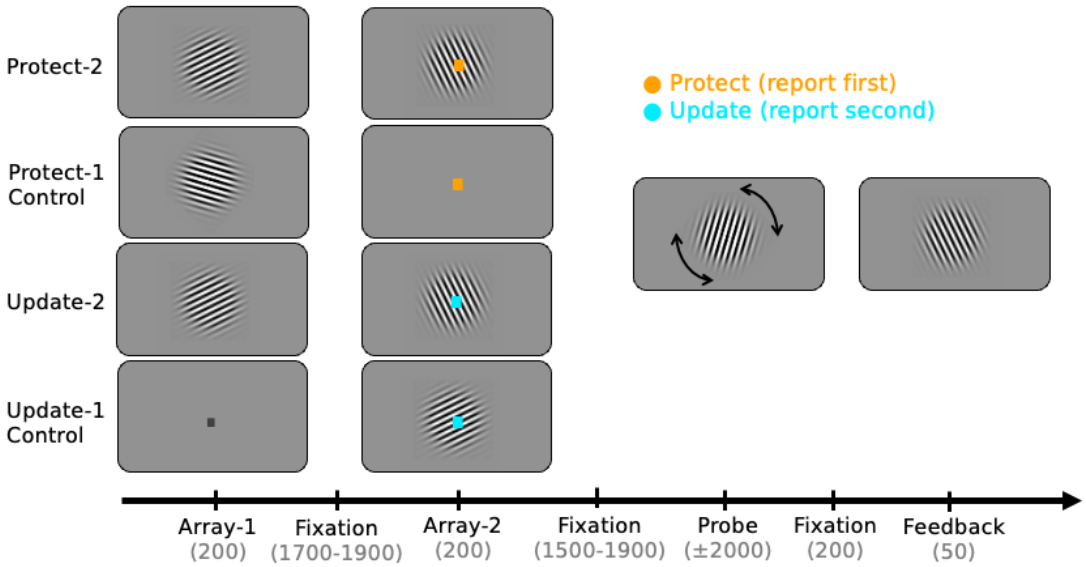

Figure 1: Experimental task design. Participants were presented with two arrays that contained an oriented bar grating or a place holder fixation dot. An oriented grating was presented either at the first (25% of trials), second (25%), or both (50%) array onset times. A colour cue that was presented on array 2 indicated the relevant stimulus for recall. The colour cued either the first array ("protect WM content") or the second array ("update WM content"). On trials where only one orientation was presented the cue was redundant but presented. Note that the colour cue contingencies are displayed for visualisation purposes; mappings between colour and cue meaning were counterbalanced. After matching the orientation in the probe display to the orientation in memory, participants were presented with feedback in the form of a grating with the correct orientation. Timings in brackets are in milliseconds. Cue size in array-2 is enlarged for visualisation purposes.

2.1 Error Rates

Participants were accurate in reproducing the target orientation (mean response error 11.73° ± 0.70° SEM; mean standard deviation 17.61° ± 1.07° S.E.M., see Table 1 for condition-wise performance). A 2-by-2 repeated-measures ANOVA showed main effects of cue type ($F_{1,19} = 16.49, p < .001, \eta^2 = 0.374$) and stimulus load ($F_{1,19} = 29.78, p < .001, \eta^2 = 0.075$). Cue type was significant for both load conditions, with absolute error higher on protect than on update trials for both load 2 ($t_{19} = 3.972, p < .001, d = 0.888$; see **Table 1**) and load 1 ($t_{19} = 3.948$, Bonferroni-corrected $p < .001, d = 0.883$). By contrast, load primarily affected the protect conditions. Error was higher on protect (2 items) than on protect (1-item control) trials ($t_{19} = 5.665, p < .001, d = 1.267$) but did not significantly differ between update (2 items) and update (1-item control) trials ($t_{19} = 1.885, p = .075, d = 0.421$), leading to a significant interaction between the two factors ($F_{1,19} = 10.90, p = .004, \eta^2 = 0.026$). Analyses using mixture modelling (Bays, Catalao, & Husain, 2009) confirmed that errors originating from responses to the non-cued grating orientation were rare (swap rate of .033 ± .01 on two-item trials; see also Huang, 2020).

2.2 Performance biases within trials

We analysed within-trial biases in behavioural performance by assessing whether the reported orientation was systematically reported as closer to or further away from the non-target orientation on the same trial (see **Methods**). **Figure 2A** shows the performance bias for all absolute angular

Table 1: **Descriptive statistics of error for all experimental conditions.**

	Protect	**Update**
1 item	11.91° (±4.61°)	8.96° (±2.52°)
2 items	14.69° (±5.38°)	9.79° (±2.89°)

Note. Mean Absolute Error (± Std. Dev.). All units are in degrees.

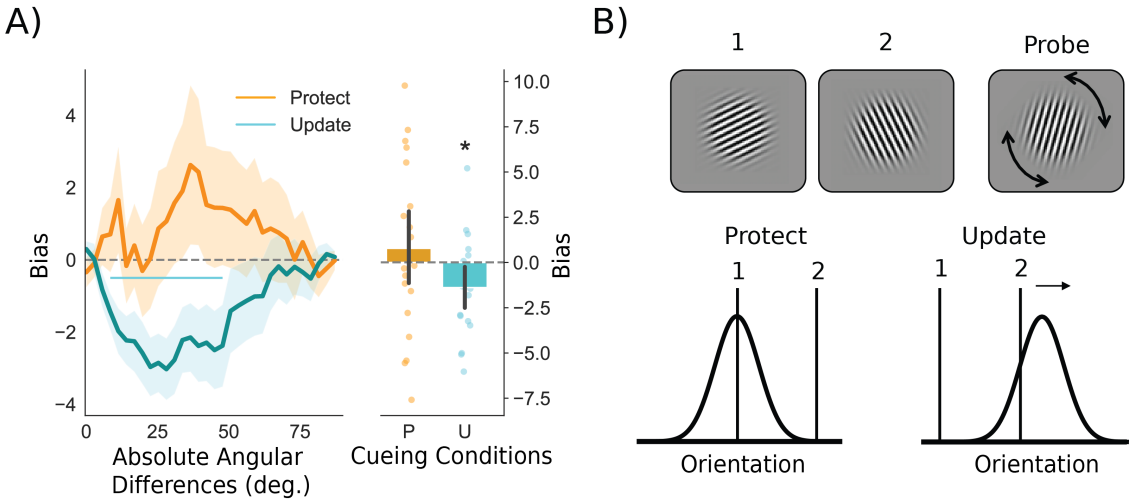

Figure 2: Within-trial repulsive performance biases. A) Within-trial bias as a function of the absolute angular distance between the two presented orientations. Trials were split for protect and update cues on load-2 trials. Shading indicates standard error of the mean. The horizontal line indicates a cluster of significant performance biases for a range of angular difference. The right panel shows the sum of integrals per subject. Error bars show 95% confidence intervals. P = protect trial; U = update trial. B) Schematic of the within-trial performance bias. A repulsive performance bias is observed for update trials only.

distances between the first and second grating orientation for update and protect trials. On trials with protect cues, there was no significant bias towards or away from the interfering second grating orientation that was not relevant to the task at hand $(t_{19} = 0.74, p = .467)$. In contrast, trials with update cues revealed significant biases away from the initially encoded first grating orientation $(t_{19} = -2.33, p = .031;$ illustrated in **Figure 2B**). The repulsive bias on update trials was confirmed using a cluster-based permutation test, showing a significant (cluster-corrected p = .012) cluster when the angular distance between the two orientations was between 8° and 48° (**Figure 2A**). Note that this analysis could not be performed on load-1 control trials, as these trials only contained a single item.

2.3 Attractive Performance Bias Between Trials

We next evaluated the between-trial bias on responses of the current trial towards the orientation that was cued on the previous trial (**Figure 3**). We assessed the performance bias as a function of angular differences between the target grating on the current and on the previous trial. The analysis also considered the position of the target grating on the current trial (1st or 2nd) and the stimulus load on the current trial (load 1 or 2). For consistency, we term all trials where participants report the first grating orientation *protect trials* and trials where participants report the second orientation *update trials*, regardless of the number of gratings presented. Again, we quantified the integral of the bias

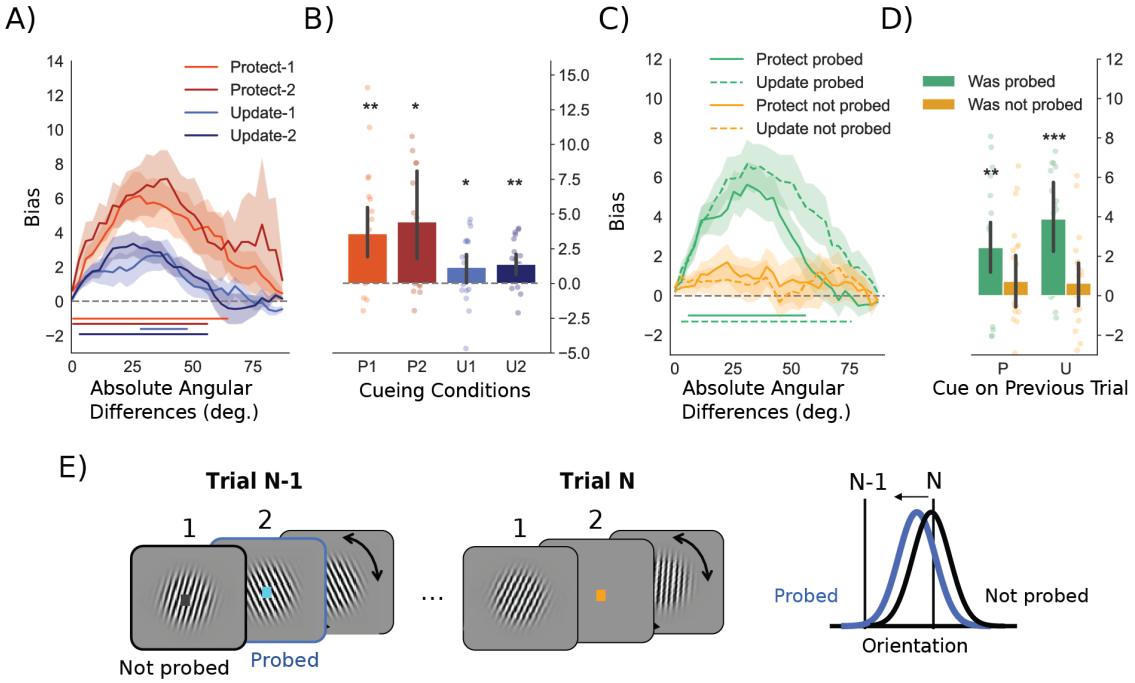

Figure 3: Attractive behavioural biases towards the relevant target orientation on the previous trial. A) Line plots for the attractive between-trial bias as a function of the absolute angular distance between the target orientation on the previous trial and the presented grating. Shading indicates standard error of the mean. Horizontal lines indicate clusters of significant performance biases for a range of angular difference. B) Bar plots indicate the average sum of integrals per subject and per condition as seen in panel 3A. C) Line plots show the attractive between-trial bias as a function of the absolute angular distance between the probed angle on the previous trial (green) or the presented but not probed angle in the previous trial (orange). The horizontal lines indicate cluster-corrected angular distances for which the bias is significant. D) Integrals of the response data shown in panel 3C where positive values indicate an attractive bias. P = Protect cue on previous trial; U = Update cue on previous trial. E) Schematic illustrating the attractive behavioural performance bias between trials. Behavioural responses on trial N were biased towards the orientation that was tested for recall (depicted in blue) and not towards the orientation that was not tested (black) on trial N-1. All error bars indicate 95% confidence intervals.

across angular distances between targets on the current and previous trial **Figure 3A, B**. In contrast to the repulsive bias described in the previous section, here we found that all conditions showed an attractive performance bias (all p < .05 in two-sided statistical tests; illustrated in **Figure 3E**). The attractive serial bias was most pronounced for small to intermediate angular distances between the inducer and current orientation (0 - 60°). A repeated-measures ANOVA on the integrals indicated an effect of cue type, with larger biases occurring in protect trials ($F_{1,19} = 5.706, p = .027, \eta^2 = .172$), but not of the number of gratings presented in a trial ($F_{1,19} = .980, p = .335, \eta^2 = .007$). The two factors did not interact ($F_{1,19} = .377, p = .547, \eta^2 = .002$). This shows that the bias was stronger when recalling the first item, which was encoded closer in time to the previous trial.

2.4 Task Dependence of Attractive Performance Bias Between Trials

Next, we repeated the same between-trial analyses but investigated the role of task relevance and cue type on the previous trial. This allowed us to test for and compare behavioural biases elicited by the probed (and reported) orientation and by the unreported orientation on the previous trial. We also tested whether the cue type on the previous trial affected bias on the current trial. We only looked at trials in which two orientations were presented on the previous trial.

A repeated-measures ANOVA on the integral of the performance bias for task relevance and cue type confirmed an effect of task relevance ($F_{1,19} = 14.684, p = .001$; **Figure 3D**) but showed no effect of cue type ($F_{1,19} = 1.423, p = .248$) or interaction ($F_{1,19} = 1.633, p = .216$). Task-relevant orientations in trials of either cue type led to a significant bias (previous trial protect (2 items): $t_{19} = 3.524, p = .002$; previous trial update (2 items): $t_{19} = 4.476, p < .001$). Unprobed orientations did not lead to a significant bias (both $p > .2$). No reliable difference was observed between the strength of the bias between update (2 items) or protect (2 items) conditions on previous trials ($t_{19} = 1.691, p = .107$). Following up, we assessed the performance bias as a function of the angular distance between the current target orientation and previously presented orientations. Cluster-based permutation testing showed an attractive bias towards the task-relevant orientation on the previous trial when a protect cue ($p = .001; 6° – 56°$; **Figure 3C**) or an update cue ($p < .001; 3° – 76°$) was presented on the previous trial. For unreported orientations no bias was observed (no candidate clusters for protect or update cue). Together, this pattern of results is summarised in **Figure 3E**, showing an attractive between-trial bias, but only with regard to items that were relevant in the previous trial.

2.5 Neural Classification of Presented Orientations

For our classification analysis, grating orientations were binned into 10 equally-spaced bins. We applied Linear Discriminant Analysis (LDA) on spatial and temporal features from all 306 MEG sensors ranging from 400 ms prior to stimulus onset up to 900 ms post stimulus onset (see **Methods**). If orientation information was present, LDA likelihood estimations gave rise to tuning curves centered on the presented orientation that could be convolved with a cosine function to result in a single evidence estimation per time point. LDA classification reflected significant evidence for the presented orientation after visual onset of the grating **Figure 4**. This revealed both significant decoding of the first grating orientation ($100 – 615$ ms; $p < .001$) as well as the second grating orientation ($95 – 590$ ms; $p < .001$). A non-significant trend was observed for the evidence of the first grating orientation following the presentation of the second grating ($250 – 600$ ms; $t_{19} = 1.971; p = .063$; see also cross-decoding analyses below). There was no difference in classifier evidence after an update or protect cue ($250 – 600$ ms; $t_{19} = 0.798; p = .435$).

2.6 Within-trial Classification Bias Away From Previous Stimulus

Tuning curves were used to estimate direction and magnitude of neural biases in orientation representation. By training the classifier on all stimuli, orientation history biases should cancel out, allowing test on trials with specific prior orientations (clockwise vs. counterclockwise to current stimulus) to reveal any neural biases. To investigate how information from the second grating was modulated by information from the first grating, we separately assessed trials with protect and

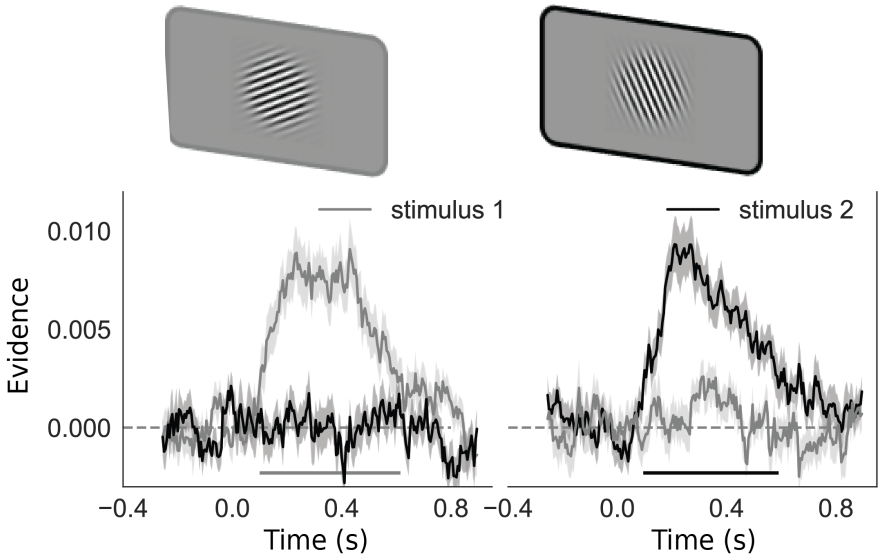

Figure 4: Orientation decoding of presented grating. Training an independent classifier to decode the first or second grating orientation in either the first- or second-time interval. The cue in the second display is removed for visualisation purposes. Error bars indicate standard error of the mean. Horizontal lines indicate periods of significant grating orientation decoding after cluster-based permutation tests against zero, p = .05.

update cues. Doing so, we evaluated the LDA evidence tuning curves elicited by the second grating. For these analyses we again only selected trials where both gratings were presented. We separated trials where the first grating was clockwise vs. counterclockwise relative to the second grating (angular distance of 10° - 50°, based on behavioural results, see **Figure 2A, B**). We considered the average of time points between 250 – 600 ms for all future analyses, since in this time window stimulus orientation could be decoded with reliable accuracy (see **Figure 4** and grey-shaded area in **Figure 5**). Echoing the performance biases, no significant bias occurred on protect trials (**Figure 5A, B**; $t_{19} = -0.723; p = .478$). However, on update trials, we observed a repulsive effect, away from the previously encoded grating orientation (**Figure 5C, D**; $t_{19} = -3.52; p = .002$).

There was no correlation, across participants, between the magnitude of the bias in the behavioural and neural data on protect trials ($r = .010; p = .968$) or update trials ($r = -.328; p = .158$).

2.7 Repulsive Neural Biases Between Trials Away From Sensory History

To probe for neural between-trial biases, we used the same approach as for within-trial neural biases. We tested the LDA-evidence derived during the stimulus encoding period for systematic deviations in likelihood estimations as a function of the angular distance between the current orientation and the probed orientation on the previous trial. If the behavioural between-trial bias reflects neural modulation during encoding of sensory features, we would expect to see an increase in likelihood for orientations presented on the previous trial, in line with the attractive performance bias. We trained the classifier on data following presentation of both the first and second grating orientation combined and included all cue types and loads. Informed by our behavioural analyses on between-trial performance biases, for the test set we selected trials where the previous probe angle had a relative difference of 0° – 60° (**Figure 2C, D**; derived from significant angular differences) CW or

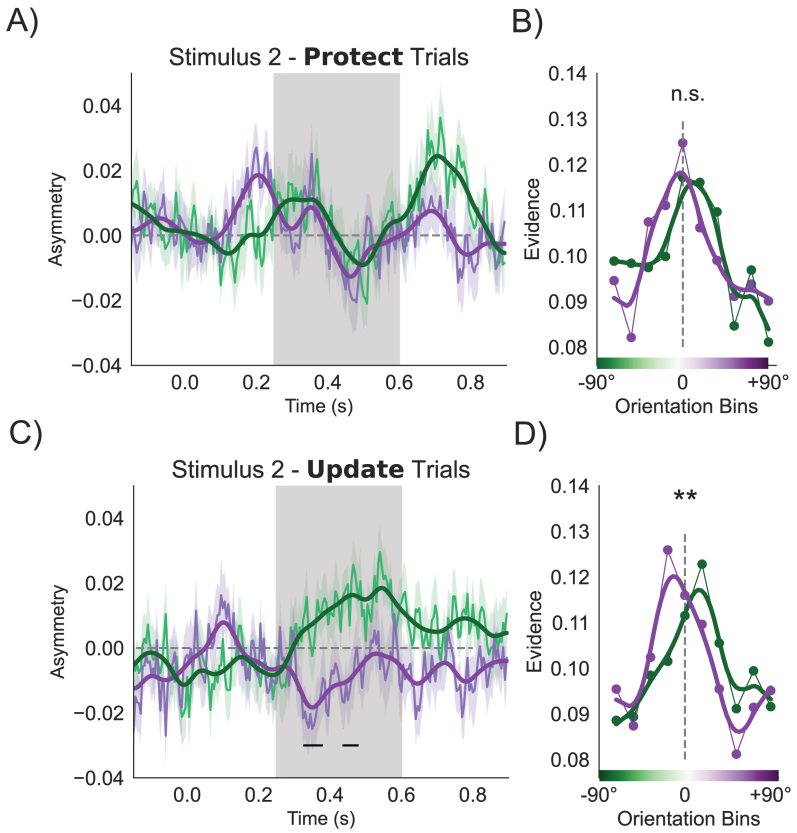

Figure 5: Shift in tuning curve relative to previous grating orientation. A) Asymmetry scores following the presentation of the second grating on protect (2 items) trials. The first grating orientation was oriented CCW (purple) or CW (green) relative to the current grating orientation. The shaded area indicates the time interval (250 - 600 ms) used for statistical analysis and to generate the tuning curves in the right panel. Smoothing (thicker lines in darker colours) was applied for visualisation purposes. B) shows the tuning curves with evidence for the presented grating orientation for trials with CW or CCW angular distances relative to the first rating. For visualisation purposes, data points were interpolated (from 10 to 50 data points) and fitted using a Savitzky-Golay filter (with a window length of 9 data points and polynomial of order 1). C) Asymmetry scores for update (2 items) trials, where participants encoded the grating orientation on screen into working memory with a bias away from the grating orientation that was presented earlier on the same trial. Shading indicates standard error of the mean. D) shows data from the grey-shaded region in panel 5C. Horizontal lines indicate periods of significant bias on unsmoothed data after cluster-based permutation tests against p = .05.

CCW from the presented grating orientation. The results were qualitatively the same and remained significant when other angular ranges were selected.

Contrary to our expectations, classifier evidence was significantly shifted away from the target orientation on the previous trial (250 – 600 ms post grating onset; $t_{19} = -2.83, p = .011$; **Figure 6A, B**) rather than mirroring the attractive behavioural bias. In practice, this would mean that if the cued orientation on the previous trial was CW, classifier evidence for CCW bins increased, and vice versa. The repulsive bias away from the target on the previous trial trended towards significance for the first ($t_{19} = -1.78, p = .090$) and was significant for the second grating ($t_{19} = -2.43, p = .025$) in the current trial when considered separately (**Figure 6C**). The between-trial repulsive bias during stimulus-two processing was present if no orientation was presented in the first interval ($t_{19} = -2.69, p = .014$) but not when the first grating was also presented ($t_{19} = -1.68, p = .110$).

Next, we tested whether this repulsive neural bias was affected by the task relevance of the grating and by the cueing condition of the previous trial. We quantified the bias for task-relevant and task-irrelevant grating orientations on the previous trial. Only trials where the previous trial was load-2 were included in this analysis. When averaging the neural bias over 250 – 600 ms, the task-relevant orientation showed a repulsive neural bias (t_{19} = -2.78, p = .012), but we found no neural bias for task-irrelevant orientations $(t_{19} = -0.33, p = .743)$, though this difference did not reach significance $(t_{19} = 1.80, p = .088)$. Yet, cluster-based permutation testing indicated a significant cluster where task-relevant orientations had a significantly stronger repulsive bias than task-irrelevant orientations $(500 – 550$ ms$; p = .039)$.

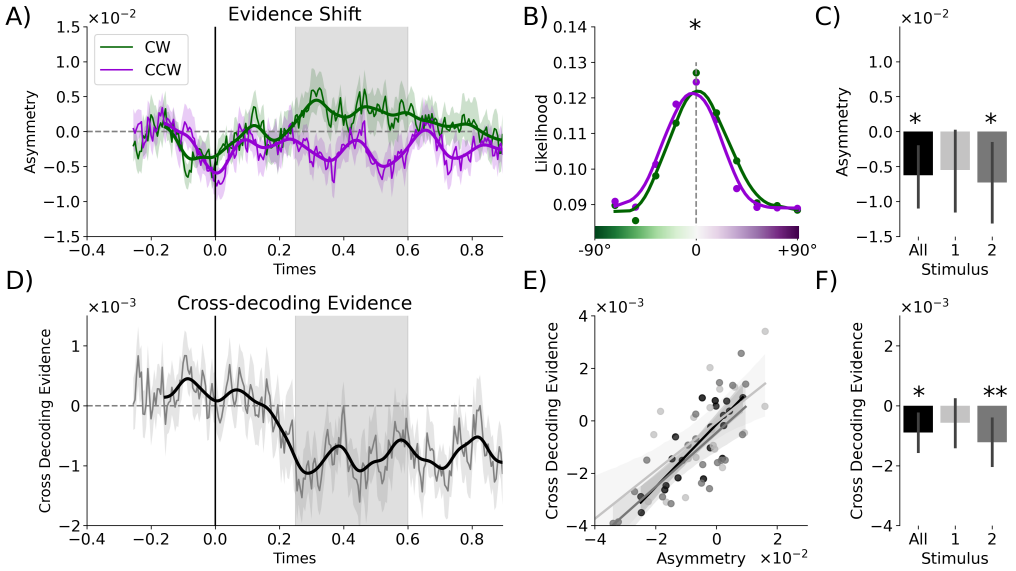

Figure 6: Bias on encoding imposed by previous trial. A) Asymmetry index for trials with a CW or CCW angular distance (0° – 60°) between the cued item on the previous trial and the presented grating (collapsed across the first and second grating). Asymmetry index time course is shown relative to the onset of grating presentation. A one-dimensional Gaussian filter with a kernel of 15 ms was applied to the data for visualisation purposes. The grey shaded area indicates which time points (250-600 ms) are used for analyses described in text and in panels 6B, C, and E. B) Average LDA orientation-likelihoods for trials with CW (green) and CCW (purple) angular distances with respect to the previous trial. Evidence is lower for CW orientation bins when in the previous trial a CW orientation was cued, and vice versa for CCW orientations. The same interpolation was applied as described in Figure Figure 5. C) Average bias relative to the previously cued orientation across all stimuli and for the first and second orientation separately. D) Cross-decoding evidence for the previously cued orientation (classifier trained on current orientation), locked to grating presentation. Same conventions as in panel A. E) Correlations between asymmetry index panel 6A and cross-decoding evidence from panel D for grating orientation processing of all stimuli combined (black), first grating (light grey) and second grating (dark grey). A least-squares linear regression model was applied for each of the conditions, with shaded 95% confidence intervals estimated using bootstrapping. F) Cross-decoding evidence from panel 6D summarised for all stimuli and separately for the first and second grating, with 95% confidence intervals. *: p < .05, **: p < .01

2.8 Cross-Decoding Evidence for Previously Presented Stimuli

If information about the previously presented orientation is still partially present in the visual system during and after the presentation of the current grating orientation, it could interact with encoding, possibly leading to the observed repulsive bias. One possibility is that the lingering representation is in an orthogonal representational format, which is different from sensory coding of features

(Libby & Buschman, 2021). In this case there would be little to no overlap between the activation pattern elicited during sensory input and the pattern related to the lingering representation of that past grating orientation. A classifier trained to separate perceptual information would therefore not cross-generalise if tested on the memory code. Alternatively, the lingering code could be present in a stable representation that shares similarity with the representation of incoming sensory information. If this were the case, a classifier, trained on the data from the current grating orientation, would cross-generalise and identify information about the past grating.

We adapted cross-decoding to detect lingering orientation-selective activity from the previous trial (also see Wan, Cai, Samaha, & Postle, 2020). After training the classifier for the presented orientation and testing for evidence of the previous trial's target orientation, we observed significant negative classifier evidence in the period of 250 – 600 after grating onset (**Figure 6F**; concatenating over grating one and grating two: $t_{19} = -2.820, p = .011$; grating one alone, $t_{19} = -1.376, p = .185$; grating two, $t_{19} = -2.951, p = .008$).

Negative classifier evidence indicates that, while information is still present about the previous orientation, orientation-selective patterns may be sign-reversed relative to stimulus encoding. This suppression of evidence for the previous trial's orientation could have been the cause of the apparent repulsive bias in the decoding of the current trial's orientation. If this was the case, we would expect the two measures to be correlated: stronger suppression of the previous orientation (negative classifier evidence) should lead to a stronger repulsive bias for the current orientation. We tested this using Pearson correlations across participants and found a significant correlation when assessing all grating presentations together (**Figure 6E**; $r = .832, p < .001$; grating one alone, $r = .685, p < .001$; grating two alone, $r = .751, p < .001$).

No correlation was observed between the attractive behavioural performance bias and the magnitude of the repulsive shift in the neural data ($r = .186, p = .432$) nor with the magnitude of negative decoding ($r = .144, p = .544$).

3 Discussion

The present study investigated biases from previously perceived and memorised information on neural coding and behavioural responses. We observed both neural and behavioural biases that demonstrate interactions between past and present sensory processing. Orientations presented on the same trial exerted repulsive biases on currently perceived orientations. In contrast, task-relevant orientations on the previous trial exerted an opposite, attractive bias, altogether providing evidence for two counteracting biasing processes. Multivariate decoding of neural data indicated that the two types of performance biases may arise from modulations acting on different stages of stimulus processing. Repulsive biases reflect at least in part modulation of sensory processing during item encoding, possibly reflecting mechanisms akin to visual adaptation that could promote visual discriminability. In contrast, the attractive performance bias observed across trials had no equivalent modulation at the sensory level. Interestingly, a neural repulsion occurred instead, suggesting the operation of post-perceptual modulatory mechanisms that can override any early sensory modulation and lead to attractive performance biases.

The present behavioural results provide evidence of two types of performance bias. Firstly, We identified a repulsive performance bias away from the first orientation in a trial. This bias was observed after participants updated their working-memory content to represent a second orientation.

There was no retrospective repulsive bias when participants ignored the second grating and reported the first orientation, even though the second grating was presented closer in time to the probe. Secondly, participants' responses on the current trial were biased towards task-relevant orientations on the previous trial (c.f., Bae & Luck, 2020; Fischer, Czoschke, et al., 2020). The attractive bias appeared strongest for items that were presented closest in time to the previous trial (first grating presentation) but was also present for the second grating, albeit with a smaller magnitude. Together, these results confirm a repulsive within-trial and an attractive between-trial performance bias.

Previously perceived stimuli generated a repulsive shift in orientation decoding, observed shortly after onset of grating presentation. This repulsive bias was only present when the presented orientation was task-relevant and therefore encoded into working memory. By contrast, task-irrelevant gratings that were not encoded into working memory could be decoded with similar precision, but did not exhibit a significant bias. Gratings from previous trials that were associated with an attractive performance bias also led to repulsive neural biases during working memory encoding. Only task-relevant orientations led to a repulsive bias. These results suggest a consistent bias that is carried forward from previous sensory processing. This repulsive bias may arise from sensory adaptation, but in the present study appears not to be a completely automatic process. Instead, the modulation by task-relevance suggests the bias is context-dependent.

Visual adaptation has been proposed as the cause of repulsive neural biases between sequentially presented stimuli that arise automatically during the early perceptual stage (Jazayeri & Movshon, 2006, 2007; Kohn, 2007; Stocker & Simoncelli, 2008; Webster, 2015) . Visual adaptation can be interpreted in terms of downweighting of neurons sensitive to the inducer orientation (Clifford, Wenderoth, & Spehar, 2000; Wainwright, 1999). Visual adaptation to previously perceived stimuli allows for efficient resource use (Stocker & Simoncelli, 2008; Webster, 2015) because neurons can code for a larger range of stimuli when their responses are not saturated. By adjusting the sensitivity levels of neurons to their context and perceptual history, saturation can be reduced, and more information can be transmitted (Webster, 2015). While sensitivity adaptations have predominantly been described in early visual areas for long-duration exposure to the inducer orientation, adaptation has been observed even after very brief exposures (see introduction) and could therefore in part explain the current results. However, visual adaptation is generally thought of as an automatic phenomenon that is not modulated by top-down contextual factors (Kohn, 2007; Webster, 2015), which is inconsistent with our observation of repulsive neural bias only for task-relevant stimuli.

We observed a significant repulsive bias only on update trials, when the presented orientation was task-relevant and encoded into working memory, but not on protect trials, when it was task-irrelevant. This task-dependent modulation was unlikely the result of reduced processing of task-irrelevant stimuli, as overall orientation decoding was not affected by task relevance. In line with recent studies, it is possible that a context-sensitive repulsive bias, possibly occurring at a post-perceptual stage (Fritsche & de Lange, 2019; Zamboni, Ledgeway, McGraw, & Schluppeck, 2016), exists alongside an early perceptual bias based on visual adaptation (Fritsche, Mostert, & de Lange, 2017). The interaction with task relevance suggests that stimulus processing is only biased when it is primed for later use in upcoming behaviour. Such an interplay of sensory and action systems has been shown to affect the encoding of visual features (Boettcher, Gresch, Nobre, & van Ede, 2021; Myers, Stokes, & Nobre, 2017; van Ede, Chekroud, Stokes, & Nobre, 2019). Circuits involved in the previous response could interact with motor encoding of the subsequent trial. An alternative cause of the task-dependent neural bias could be that participants compare the relative orientation of the two gratings involuntarily when they are presented in sequence, generating a repulsive bias

(Bae & Luck, 2017; Chunharas, Rademaker, Brady, & Serences, 2019; Czoschke, Fischer, Beitner, Kaiser, & Bledowski, 2019; Czoschke, Peters, Rahm, Kaiser, & Bledowski, 2020). The repulsion could segregate the orientations for better storage of individualised items (e.g. Wei, Wang, & Wang, 2012). Under this explanation, only attended and encoded features would be subject to interactions with previously or presently stored features.

We observed a marked contrast between the attractive performance bias towards the orientation of the previous trial and the repulsive neural bias during encoding. This result appears to be at odds with previous studies that have assigned an early perceptual origin to attractive between-trial biases observed in behaviour (e.g., Cicchini, Mikellidou, & Burr, 2017; Cicchini, Mikellidou, & Burr, 2018; Fischer & Whitney, 2014). By contrast, we failed to find evidence of an attractive modulation of sensory representations towards the orientation of the previous trial at the working memory encoding stage. There have been few studies that have studied brain activity directly and have investigated shifts in neural - and behavioural - representations as a function of sensory history. EEG studies have shown that previous trial information can be decoded during the encoding phase of the current trial (Bae & Luck, 2019) or immediately prior to the current trial (Barbosa & Compte, 2020), and visually evoked neural responses in numerosity judgement tasks are modulated by stimulus history (Fornaciai & Park, 2018, 2019). These results showing lingering effects of stimulus history have been interpreted as the neural basis of the attractive performance bias. But the mere presence of prior stimulus information does not indicate how it will influence processing of the current stimulus. St. John-Saaltink et al. 2016 observed an attractive behavioural bias and a neural bias in early visual areas using fMRI. However, this study used only two stimulus orientations (45° and 135°) whose offset is too large to produce a reliable behavioural bias in other settings, making it difficult to relate their findings to the present results. A more recent fMRI study (Sheehan & Serences, 2021) instead found, across several visual areas, repulsive neural biases relative to the orientation on the previous trial, in spite of an attractive behavioural bias. These results are in line with the present findings and together make a case that prior stimuli lead to repulsion at the encoding stage, and that attractive performance biases arise elsewhere. The poor temporal resolution of the BOLD response can make it difficult to disambiguate early visual responses during encoding from re-entrant post-perceptual processes influencing responses in sensory areas. Our results further indicate that the repulsive bias arises within 500 ms of stimulus onset, and that they occur simultaneously relative to multiple prior stimuli (from the same trial and the previous trial). Finally, single-unit recordings from frontal eye fields (FEF) have demonstrated a repulsive neural bias, paired with an attractive behavioural bias (Papadimitriou, White, & Snyder, 2017). Our results therefore suggest that neural repulsive biases observed previously in frontal cortex may generalise to parts of the brain involved in working memory encoding.

Critically, the present study investigated sensory processing to infer the directionality of the resulting bias (Wolff, Jochim, Akyürek, Buschman, & Stokes, 2020). We tested this using cross-decoding analyses - that train on the presented orientation and predict previous orientations – to show negative or inverted evidence for the stimuli from the previous trial. Congruent with a suppressive bias, patterns related to the orientation presented on the previous trial are downweighted, which implies a relative increase in the pattern strength for all other orientations. In turn, this may shift the tuning curve related to the current orientation away from the inducer, generating a repulsive bias. This relationship was confirmed by the robust correlation between cross-decoding and repulsive bias magnitude.

Altogether, through the neural bias analyses in this study, we demonstrate a consistent repulsive shift in neural evidence during working memory encoding. Our results imply that perceptual adaptation, along with context-sensitive factors, contributes to feature-selective downweighting to exert a repulsive bias away from recent stimulus features. Interestingly, no evidence of an attractive neural bias acting directly on sensory aspects of encoding was observed. Neural data thereby provide indirect evidence for the post-perceptual account of attractive between-trial biases, rather than modulating encoding stages (Bae & Luck, 2020; Bliss, Sun, & D'Esposito, 2017; Fritsche, Mostert, & de Lange, 2017; Kim, Burr, Cicchini, & Alais, 2020; Pascucci et al., 2019). We speculate that the attractive between-trial bias instead arises through post-perceptual processing stages involving memory (Bliss, Sun, & D'Esposito, 2017; Fritsche, Mostert, & de Lange, 2017), perceptual decision-making, or motor planning (Boettcher, Gresch, Nobre, & van Ede, 2021; Machado, Neto, Bartels, Israelita, & Einstein, 2021; Sadil, Cowell, & Huber, 2021) . The source of the attractive between-trial bias, whatever its neural mechanism, may be strong enough to overcome the repulsive bias that we see during the perceptual/encoding stage. Together, these co-existing biases may help guide efficient coding for nuanced perceptual discriminations and visual stability across our environment.

4 Methods

4.1 Participants

Twenty healthy volunteers with normal or corrected-to-normal vision participated in the study. All participants were between 20 and 36 years old (mean 25.4 years old; eleven females). Prior to taking part in the study, volunteers provided their informed consent according to the procedures approved by the Central University Research Ethics Committee of the University of Oxford. Participants received £15 per hour compensation for taking part in this study.

4.2 Experimental set-up

Participants sat in the MEG scanner, which was situated in a dimly lit, sound-proof, and magnetically shielded room. A projection screen was placed at a viewing distance of 90 cm. Visual stimuli were projected at the back of the screen at a spatial resolution of 1024×768 pixels using a refresh rate of 60 Hz using a Panasonic DLP projector (PT-D7700E).

The task was programmed and presented using Matlab (Mathworks, Nantick, WA) in conjunction with the Psychophysics Toolbox (Brainard, 1997). Participants indicated their responses on an optic-fibre response box.

4.3 Task

Participants performed a precision working-memory task in which they reproduced the orientation of one of two grating stimuli presented sequentially with independent orientations (**Figure 1**). Simultaneously with the presentation of the second grating, participants were cued as to which grating orientation to report when probed at the end of the trial. On half of the trials only the first or second grating was presented and used for reporting.

Each trial started with a central fixation cross (0.2° visual angle) on screen for 800 ms with a grey background (RGB: 127, 127, 127; **Figure 1**). Subsequently, a sinusoidal Gabor stimulus at a random angle was centrally presented on the visual display for 200 ms (diameter of 6° visual angle, 2 cycles per degree of visual angle, 50% contrast, tapered by a Gaussian envelope with a 1.5° standard deviation). On trials where the first grating was not presented the fixation dot changed colour from black to grey (RGB: 192, 192, 192) to signal the omission. After a delay of 1700 – 1900 ms, the second grating was presented. This second grating had the same properties as the first grating except for the angle, which was again randomly drawn independently from that of the first grating orientation. At the centre of the grating, the fixation dot changed colour (orange: (255, 161, 0); cyan: (0, 236, 255)) to signal that either the first orientation (*protect* working memory content) or that the second orientation would be probed (*update* working memory content).

Cue-colour contingencies of the experiment were counterbalanced across participants and changed halfway through the experiment. After the colour contingencies switched, participants practised the new contingencies over the course of one block before continuing with the second half of the session. The colour cue was always valid in indicating which grating orientation was relevant for the reporting stage at the end of the trial. On trials where the second grating was absent the cue would therefore signal the relevance of the first array.

After another delay of 1700 – 1900 ms, a probe grating was presented, which participants were asked to adjust to match the cued grating in memory using the response box. Adjustments were made by pressing buttons with the right hand to rotate the grating either clockwise (middle finger) or anticlockwise (index finger). Responses were confirmed by pressing a button with their left index finger. A 200-ms fixation period followed, after which participants were presented with 50 ms of feedback in the form of a grating indicating the correct orientation.

In total, participants completed 400 trials; 200 trials where two items were presented (100 with update cue and 100 with protect cue) and 200 trials where one item was presented (100 with only grating 1 and 100 only grating 2). The resulting factorial design included four conditions: protect with two items presented (protect, 2 items), protect with one item presented (protect, 1-item control), update with two items presented (update, 2 items), and update with one item presented (update, 1-item control). These were randomly mixed and presented in blocks of 50 trials, with each block lasting approximately 10 minutes.

4.4 Behavioural Analysis

Response error was quantified by calculating the circular distance of the recalled orientation to the cued orientation. All responses were mapped onto a -180° to 180° space to compute circular error. In all between-trial analyses, we excluded the first trial of each block.

4.5 Mixture Modelling

To investigate the relative rate of responses to targets, guesses, and erroneous responses to the wrong target ('swap errors'), we fit a classical mixture model frequently used in working memory literature (Bays, Catalao, & Husain, 2009; Schneegans & Bays, 2017; Zhang & Luck, 2008). The model was fit separately for each condition (protect, 2 items; protect, 1-item control; update, 2 items; update, 1-item control). The mixture model estimated the precision of the von Mises distribution, target response rate, guess rate, and swap rate to the item that was presented on the same trial but not cued (only load-2 trials). The model provided single trial estimates of each of the three parameters.

4.6 Performance Bias Calculation

We calculated the performance bias (the circular difference between the response and the target orientation) as a function of the circular difference between the target orientation and the orientation that induced the bias (the previous target orientation on the preceding trial or the task-irrelevant grating on the same trial). We computed the difference between the target orientation and inducer through subtraction, with all angular differences mapped between -90° and 90°. These distances were binned into 64 equally-sized bins. We computed the average signed error (bias) across all trials within each bin. To reduce noise, we applied a sliding window, computing the average response error over 25% of the bins (16 bins). Edge artifacts were avoided by wrapping the angular differences around, to ensure the average was always computed over 16 bins. Subsequently, we sign-flipped the bias in bins with a negative distance and averaged over negative (-90° to 0°) and positive distances (0° to 90°) between target orientation and inducing orientation, resulting in 32 bins. The integral over absolute distances (0° to 90°) calculated over the 32 bins for each condition and participant served as a measure of bias.

4.7 MEG Acquisition

Participants were seated in the MEG scanner after being instructed about the task specifics. They completed one practice block while seated in the scanner prior to MEG recording onset. Participants were instructed to maintain their gaze at the central fixation cross and to minimise blinking throughout the trial.

Neuromagnetic data were acquired using a whole-head VectorView system including 204 planar gradiometers and 102 magnetometers (Elekta Neuromag Oy, Helsinki, Finland) in a magnetically shielded room. Throughout the experiment, participants' head position was monitored continuously using index coils placed at four points on the head. Magnetic field strength was sampled at a rate of 1000 Hz and band-pass filtered on-line between 0.03 Hz and 300 Hz. In addition, vertical and horizontal electro-oculograms were measured using electrodes placed above, below, and adjacent to the eyes. Eye movements were monitored using an EyeLink 1000 (SR Research) eye tracker at a frequency of 1000 Hz.

4.8 MEG Data Preprocessing

The data were pre-processed offline using Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011), OHBA software library (OSL) drawing on SPM8 (http://www.fil.ion.ucl.ac.uk/spm), and Elekta software. Prior to any preprocessing, the MEG data were visually inspected to remove and interpolate any sensors that displayed excessive levels of noise and were subsequently de-noised and motion corrected using Maxfilter Signal Space Separation (Taulu, Kajola, & Simola, 2004) before removing independent components related to cardiac and eye-blink artefacts. Data were epoched around the first grating and second grating (from 400 ms prior to grating onset to 900 ms after onset) and downsampled to 200 Hz. Trials with high variance in either gradiometers or magnetometers were identified and excluded using a generalised ESD (extreme studentised deviate; Rosner, 1983) test at a 0.05 significance threshold. This resulted in 7.49% ± 11.55% (mean ± standard deviation) being excluded during preprocessing.

4.9 LDA Classification

Data were further pre-processed, where magnitudes of magnetometers were approximately matched to gradiometers by multiplication (factor 20) and subjected to spatio-temporal decoding as described in (Hajonides, Nobre, van Ede, & Stokes, 2021; Wolff, Jochim, Akyürek, Buschman, & Stokes, 2020; Wolff, Jochim, Akyürek, & Stokes, 2017). Data from all 306 MEG sensors across a sliding window of 30 time points (150 ms) were concatenated into a vector. Pre-stimulus baselining was not applied to maintain stable information from previously presented stimuli. Dimensionality was reduced for each time point through a principal component analysis, maintaining 90% of the variance (between 250 to 600 ms this was around 209 ± 39 components, mean ± std.). To train an LDA classifier, the data were split in train and test sets using 10-fold stratified cross-validation. Grating angles were binned into 10 equally spaced orientation bins (0° to 18°, 18° to 36°, 36° degree to 126°, 126° to 144°, 144° to 162°, 162° to 180°). For each trial and time point, we thus obtained 10 LDA distances estimating the likelihood for each of the bins. Tuning curves were constructed by aligning evidence across trials around the same category bin. In cross-decoding analyses, LDA classifiers were trained on orientation bins of one event (e.g., presented grating) but classifier evidence aligned around bins of another orientation (e.g., target orientation on the previous trial). The resulting tuning curves were convolved with a cosine.

4.10 Bias Computation

For within-trial biases, we assessed processing of the second grating and only considered load-2 trials. The classifier was trained on all presentations of the second grating and bin likelihoods were generated for each trial. For between-trial analyses, we analysed orientation processing of both the first and second grating. For this reason, we trained the classifier on all trials and generated bin predictions for all trials.

Subsequently, based on the results from the performance-bias analyses, we selected trials where the angular distance between the inducer and the grating orientation on the display led to a significant behavioural bias at the group level. In the case of the within-trial repulsive bias, the inducer was the orientation of the first grating on the same trial; for the between-trial analyses the inducer was the target orientation on the previous trial. As a dependent variable, we considered likelihood estimations for each orientation bin, where we expect the highest likelihood for the angular bin that has zero offset to the presented orientation and decreasing likelihoods for bins with larger angular distances to the presented orientation. We separately assessed likelihood estimations on trials where the inducer orientation was clockwise (CW) or counterclockwise (CCW) with respect to the current orientation. For both CW and CCW trials we separately averaged the evidence from the orientation bins CW (-72° to -18°) and evidence from the CCW bins (18° to 72°). Symmetry scores were computed by obtaining the difference between the two groups of angular bins (CW minus CCW). Finally, we calculated an overall neural bias score by subtracting symmetry scores on trials with CW vs. CCW inducers. Attractive biases resulted in a positive score (i.e., trials with CW angular distances resulted in more CW

evidence, CCW angular distances resulted in more CCW evidence), whereas repulsive biases resulted in a negative score (i.e., CW angular distances resulted in more less CW evidence than CCW trials, and vice versa).

4.11 Statistical Testing

Statistical tests were computed using both JASP (Team, 2020) and Scipy (Virtanen et al., 2020).

We tested the time series of cosine-convolved classifier evidence against zero using a cluster-based permutation test (using MNE; Gramfort et al., 2013). We ran 100,000 iterations and a F-statistic threshold of 2.093 ($p < .05$ with 19 degrees of freedom). The clusters where groups of time points were significantly different from zero are indicated in the relevant figures using horizontal lines. Cluster-based permutation testing was also applied to performance bias across angular distance between the presented orientation and the inducer orientation.

Estimating the shifts in the tuning curves relied on the presence of a tuning curve. Hence, we only used the average of all timepoints where cosine-convolved classifier evidence for the presented grating orientation was significantly above zero in our comparison analyses (in all reported time averages, time points between 250 and 600 ms were used).

All tests were two-sided unless stated otherwise.

Acknowledgements

This research was funded by an ESRC Grand Union studentship and the Scatcherd European Scholarship awarded to **J.E.H.**, an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) to **F.v.E.**, was supported by a James S. McDonnell Foundation Scholar Award (220020405) and an ESRC grant (ES/S015477/1) to **M.G.S.**, a James S. McDonnell Foundation Understanding Human Cognition Collaborative Award (number 220020448) and a Wellcome Trust Senior Investigator Award (104571/Z/14/Z) to **A.C.N.**, as well as a Wellcome Trust award (201409/ Z/16/Z) and with support from University College Oxford to **N.E.M.** The work was enabled by the NIHR Oxford Health Biomedical Research Centre and the Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

Bae, G. Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception, and Psychophysics*, *79*(8), 2376–2395. https://doi.org/10.3758/s13414-017-1404-8

Bae, G. Y., & Luck, S. J. (2019). Reactivation of Previous Experiences in a Working Memory Task. *Psychological Science*, *30*(4), 587–595. https://doi.org/10.1177/0956797619830398

Bae, G. Y., & Luck, S. J. (2020). Serial dependence in vision: Merely encoding the previous-trial target is not enough. *Psychonomic Bulletin and Review*, *27*(2), 293–300. https://doi.org/10.3758/s13423-019-01678-7

Barbosa, J., & Compte, A. (2020). Build-up of serial dependence in color working memory. *Scientific Reports*, *10*(1), 1–7. https://doi.org/10.1038/s41598-020-67861-2

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, *9*(10), 7.1–11. https://doi.org/10.1167/9.10.7

Bliss, D. P., Sun, J. J., & D'Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, *7*(1), 1–13. https://doi.org/10.1038/s41598-017-15199-7

Boettcher, S. E., Gresch, D., Nobre, A. C., & van Ede, F. (2021). Output planning at the input stage in visual working memory. *Science Advances*, *7*(13), 1–15. https://doi.org/10.1126/sciadv.abe8212

Born, R. T., & Tootell, R. B. H. (1992). Middle temporal visual area. *Nature*, *357*(June), 497–499.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/10.1163/156856897X00357

Chunharas, C., Rademaker, R., Brady, T., & Serences, J. (2019). Adaptive memory distortion in visual working memory. (743941). https://doi.org/10.31234/osf.io/e3m5a

Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(21), 7867–7872. https://doi.org/10.1073/pnas.1402785111

Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, *17*(14), 1–9. https://doi.org/10.1167/17.14.6

Cicchini, G. M., Mikellidou, K., & Burr, D. C. (2018). The functional role of serial dependence. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1890). https://doi.org/10.1098/rspb.2018.1722

Clifford, C. W., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1454), 1705–1710. https://doi.org/10.1098/rspb.2000.1198

Czoschke, S., Fischer, C., Beitner, J., Kaiser, J., & Bledowski, C. (2019). Two types of serial dependence in visual working memory. *British Journal of Psychology*, *110*(2), 256–267. https://doi.org/10.1111/bjop.12349

Czoschke, S., Peters, B., Rahm, B., Kaiser, J., & Bledowski, C. (2020). Visual objects interact differently during encoding and memory maintenance. *Attention, Perception, and Psychophysics*, *82*(3), 1241–1257. https://doi.org/10.3758/s13414-019-01861-x

Dong, D., & Atick, J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems*, *6*(3), 345–358. https://doi.org/10.1088/0954-898X/6/3/003

Fischer, C., Czoschke, S., Peters, B., Rahm, B., Kaiser, J., & Bledowski, C. (2020). Context information supports serial dependence of multiple visual objects across memory episodes. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15874-w

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Current Biology*, *16*(5), 738–743. https://doi.org/doi:10.1038/nn.3689

Fisher, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature neuroscience*, *17*(5), 738–743. https://doi.org/10.1161/CIRCRESAHA.116.303790.The

Fornaciai, M., & Park, J. (2018). Attractive Serial Dependence in the Absence of an Explicit Task. *Psychological Science*, *29*(3), 437–446. https://doi.org/10.1177/0956797617737385

Fornaciai, M., & Park, J. (2019). Neural dynamics of serial dependence in numerosity perception. *Journal of Cognitive Neuroscience*, *32*(1), 141–154. https://doi.org/10.1162/jocn_a_01474

Fritsche, M., & de Lange, F. P. (2019). The role of feature-based attention in visual serial dependence. *bioRxiv*, *19*, 1–13. https://doi.org/10.1101/584789

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite Effects of Recent History on Perception and Decision. *Current Biology*, *27*(4), 590–595. https://doi.org/10.1016/j.cub.2017.01.006

Fritsche, M., Solomon, S. G., & De Lange, F. P. (2021). Brief stimuli cast a long-term trace in visual cortex. *bioRxiv*, 2021.02.10.430579. https://doi.org/10.1101/2021.02.10.430579

Fritsche, M., Spaak, E., & de Lange, F. P. (2020). A bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *eLife*, *9*, 1–32. https://doi.org/10.7554/eLife.55389

Gekas, N., McDermott, K. C., & Mamassian, P. (2019). Disambiguating serial effects of multiple timescales. *Journal of Vision*, *19*(6), 1–14. https://doi.org/10.1167/19.6.24

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*(267), 1–13. https://doi.org/10.3389/fnins.2013.00267

Hajonides, J. E., Nobre, A. C., van Ede, F., & Stokes, M. G. (2021). Decoding visual colour from scalp electroencephalography measurements. *NeuroImage*, *237*(April), 118030. https://doi.org/10.1016/j.neuroimage.2021.118030

Huang, J., & Sekuler, R. (2010). Distortions in recall from visual memory: Two classes of attractors at work. *Journal of Vision*, *10*(2), 1–27. https://doi.org/10.1167/10.2.24

Huang, L. (2020). Distinguishing target biases and strategic guesses in visual working memory. *Attention, Perception, and Psychophysics*, *82*(3), 1258–1270. https://doi.org/10.3758/s13414-019-01913-2

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*(5), 690–696. https://doi.org/10.1038/nn1691

Jazayeri, M., & Movshon, J. A. (2007). A new perceptual illusion reveals mechanisms of sensory decoding. *Nature*, *446*(7138), 912–915. https://doi.org/10.1038/nature05739

Kim, S., Burr, D., Cicchini, G. M., & Alais, D. (2020). Serial dependence in perception requires conscious awareness. *Current Biology*, *30*(6), R257–R258. https://doi.org/10.1016/j.cub.2020.02.008

Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial Dependence across Perception, Attention, and Memory. https://doi.org/10.1016/j.tics.2017.04.011

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, *97*(5), 3155–3164. https://doi.org/10.1152/jn.00086.2007

Libby, A., & Buschman, T. J. (2021). Rotational dynamics reduce interference between sensory and memory representations. *Nature Neuroscience*, *24*(5), 715–726. https://doi.org/10.1038/s41593-021-00821-9

Machado, R., Neto, D. A., Bartels, A., Israelita, H., & Einstein, A. (2021). Disrupting short-term memory in premotor but not visual cortex affects serial dependence in visuomotor integration. *bioRxiv*, *307*, 1–52.

Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory and Cognition*, *36*(1), 43–52. https://doi.org/10.3758/MC.36.1.43

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, *21*(6), 449–461. https://doi.org/10.1016/j.tics.2017.03.010

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/156869

Papadimitriou, C., White, R. L., & Snyder, L. H. (2017). Ghosts in the Machine II: Neural correlates of memory interference from the previous trial. *Cerebral Cortex*, *27*(4), 2513–2527. https://doi.org/10.1093/cercor/bhw106

Pascucci, D., Mancuso, G., Santandrea, E., Libera, C. D., Plomp, G., & Chelazzi, L. (2019). *Laws of concatenated perception: Vision goes for novelty, decisions for perseverance* (Vol. 17). https://doi.org/10.1371/journal.pbio.3000144

Patterson, C. A., Wissig, S. C., & Kohn, A. (2013). Distinct effects of brief and prolonged adaptation on orientation tuning in primary visual cortex. *Journal of Neuroscience*, *33*(2), 532–543. https://doi.org/10.1523/JNEUROSCI.3345-12.2013

Priebe, N. J., & Lisberger, S. G. (2002). Constraints on the source of short-term motion adaptation in macaque area MT. II. Tuning of neural circuit mechanisms. *Journal of Neurophysiology*, *88*(1), 370–382. https://doi.org/10.1152/jn.2002.88.1.370

Rosner, B. (1983). Percentage Outlier Points for Generalized ESD Many- Procedure. *Technometrics, 25*(2), 165–172.

Sadil, P., Cowell, R. A., & Huber, D. E. (2021). The yin-yang of serial dependence effects: every response is both an attraction to the prior response and a repulsion from the prior stimulus. *PsyArXiv*.

Schneegans, S., & Bays, P. M. (2017). Neural Architecture for Feature Binding in Visual Working Memory. *The Journal of Neuroscience*, *37*(14), 3913–3925. https://doi.org/10.1523/JNEUROSCI.3493-16.2017

Sheehan, T. C., & Serences, J. T. (2021). Sensory readout accounts for adaptation. *bioRxiv*, 2021.04.06.438664. https://doi.org/10.1101/2021.04.06.438664

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216. http://www.annualreviews.org/doi/pdf/10.1146/annurev.neuro.24.1.1193

St John-Saaltink, E., Kok, P., Lau, H. C., & de Lange, F. P. (2016). Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *36*(23), 6186–92. https://doi.org/10.1523/JNEUROSCI.4390-15.2016

Stocker, A. A., & Simoncelli, E. P. (2008). A Bayesian model of conditioned perception. *Advances in Neural Information Processing Systems 20 - Proceedings of the 2007 Conference*, *20*, 1409–1416.

Störmer, V. S., & Alvarez, G. A. (2014). Feature-based attention elicits surround suppression in feature space. *Current Biology*, *24*(17), 1985–1988. https://doi.org/10.1016/j.cub.2014.07.030

Suárez-Pinilla, M., Seth, A. K., & Roseboom, W. (2018). Serial dependence in the perception of visual variance. *Journal of Vision*, *18*(7), 1–24. https://doi.org/10.1167/18.7.4

Taulu, S., Kajola, M., & Simola, J. (2004). Suppression of interference and artifacts by the signal space separation method. *Brain Topography*, *16*(4), 269–275. https://doi.org/10.1023/B:BRAT.0000032864.93890.f9

Team, J. (2020). JASP (Version 0.14.1)[Computer software].

van Bergen, R. S., & Jehee, J. F. (2019). Probabilistic Representation in Human Visual Cortex Reflects Uncertainty in Serial Decisions. *Journal of Neuroscience*, *39*(41), 8164–8176. https://doi.org/10.1523/JNEUROSCI.3212-18.2019

van Ede, F., Chekroud, S. R., Stokes, M. G., & Nobre, A. C. (2019). Concurrent visual and motor selection during visual working memory guided action. *Nature Neuroscience*, *22*(3), 477–483. https://doi.org/10.1038/s41593-018-0335-6

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R., Jones, E., Kern, R., Larson, E., . . . Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Research*, *39*(23), 3960–3974. https://doi.org/10.1016/S0042-6989(99)00101-7

Wan, Q., Cai, Y., Samaha, J., & Postle, B. R. (2020). Tracking stimulus representation across a 2-back visual working memory task: Tracking 2-back representation. *Royal Society Open Science*, *7*(8). https://doi.org/10.1098/rsos.190228rsos190228

Webster, M. A. (2015). Visual Adaptation. *Annual Review of Vision Science*, *1*, 547–567. https://doi.org/10.1146/annurev-vision-082114-035509

Wei, Z., Wang, X. J., & Wang, D. H. (2012). From distributed resources to limited slots in multiple-item working memory: A spiking network model with normalization. *Journal of Neuroscience*, *32*(33), 11228–11240. https://doi.org/10.1523/JNEUROSCI.0735-12.2012

Wolff, M. J., Jochim, J., Akyürek, E. G., Buschman, T. J., & Stokes, M. G. (2020). Drifting codes within a stable coding scheme for working memory. *PLoS biology*, *18*(3), e3000625. https://doi.org/10.1371/journal.pbio.3000625

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. https://doi.org/10.1038/nn.4546

Zamboni, E., Ledgeway, T., McGraw, P. V., & Schluppeck, D. (2016). Do perceptual biases emerge early or late in visual processing? Decision-biases in motion perception. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1833). https://doi.org/10.1098/rspb.2016.0263

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–5. https://doi.org/10.1038/nature06860